

# Identification of catalytically important residues of the carotenoid 1,2-hydratases from *Rubrivivax gelatinosus* and *Thiocapsa roseopersicina*

Aida Hiseni<sup>1</sup> · Linda G. Otten<sup>1</sup> · Isabel W. C. E. Arends<sup>1</sup>

Received: 1 May 2015 / Revised: 1 September 2015 / Accepted: 10 September 2015 / Published online: 19 October 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** Carotenoid 1,2-hydratases (CrtC) catalyze the selective addition of water to an isolated carbon–carbon double bond. Although their involvement in the carotenoid biosynthetic pathway is well understood, little is known about the mechanism by which these hydratases transform carotenoids such as lycopene into the corresponding hydroxyl compounds. Key residues were identified at positions His239, Trp241, Tyr266, and Asp268 in CrtC from *Rubrivivax gelatinosus* (and corresponding positions in *Thiocapsa roseopersicina*). Alanine mutants at these positions were found to be completely inactive, suggesting their direct involvement in the catalytic reaction. Our resulting mechanistic hypothesis is in analogy with the recently studied class of terpenoid cyclase enzymes containing a highly acidic aspartic residue in their active site. We propose that a similar aspartic acid residue, which is conserved through all putative CrtCs, is involved in initial protonation of the double bond in lycopene.

**Keywords** Carotenoid 1,2-hydratase · CrtC · Hydro-lyase · Enzyme mechanism · Biocatalysis

**Electronic supplementary material** The online version of this article (doi:10.1007/s00253-015-6998-y) contains supplementary material, which is available to authorized users.

✉ Isabel W. C. E. Arends  
I.W.C.E.Arends@tudelft.nl; <http://www.bt.tudelft.nl/boc>

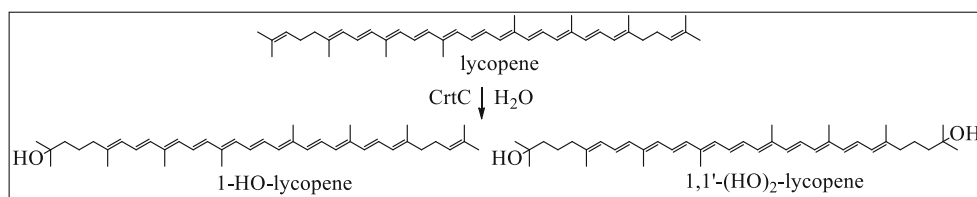
<sup>1</sup> Biocatalysis Group, Department of Biotechnology, Delft University of Technology, Julianalaan 67, 2628 BC Delft, The Netherlands

## Introduction

Carotenoids, which represent one of the most abundant natural pigments with structural and protective properties (Armstrong and Hearst 1996), play an essential role in the photosynthetic machinery of phototrophic organisms such as purple bacteria (Jensen et al. 1961) and higher plants (Cazzonelli 2011). In addition, they have been identified in fungi and some non-photosynthetic bacteria (Armstrong 1997). Carotenoid 1,2-hydratase (also known as CrtC) is a member of hydro-lyase group EC 4.2.1.131 (Hiseni et al. 2015). The enzyme takes part in the biosynthetic pathway of carotenoids (Umeno et al. 2005). CrtC introduces a tertiary hydroxyl group into an acyclic carotenoid molecule by addition of water to the carbon–carbon double bond at the C-1 position. The enzyme belongs to Pfam family PF07143 that encompasses members from several purple photosynthetic bacteria. On the other hand, CrtCs have been identified, which are able to hydrate monocyclic carotenoid gamma-carotene. These are evolutionary very distinct from the PF07143 members and have been given the name CruF (Sun et al. 2009).

To date, two representatives of the PF07143 family, the CrtCs from purple non-sulfur Betaproteobacteria *Rubrivivax gelatinosus* and purple sulfur Gammaproteobacteria *Thiocapsa roseopersicina*, were recombinantly expressed and characterized (Hiseni et al. 2011). Biochemical studies have revealed that these enzymes are able to convert cofactor-independently lycopene into 1-HO-lycopene and 1,1'-(HO)<sub>2</sub>-lycopene (Fig. 1). In addition, they showed some activity towards the unnatural substrate geranylgeraniol, a C20 molecule that resembles the natural substrate lycopene.

CrtCs are appealing enzymes in the biotechnology field because they are able to generate a tertiary alcohol, a highly valuable building block for the synthesis of several bioactive natural products and pharmaceuticals (Hiseni et al. 2015; Kourist et al.



**Fig. 1** Reaction catalyzed by *Rubrivivax gelatinosus* and *Thiocapsa roseopersicina* carotenoid 1,2-hydratase; the conversion of lycopene into 1-HO-lycopene and 1,1'-(HO)<sub>2</sub>-lycopene

2008). Furthermore, they possess an intrinsically high stability at a wide pH and temperature range, which constitute useful properties for an industrial application (Hiseni et al. 2011). The subcellular location of this enzyme in the cell membrane fraction (membrane associated) allows for a straightforward isolation and simplified large-scale purification.

From a chemical point of view, CrtCs are able to perform a challenging chemical reaction, namely the selective addition of water to an isolated carbon–carbon double bond (Jin and Hanefeld 2011). Using the enzyme, the reaction proceeds without assistance of electron-withdrawing groups or transition metal cations, while the chemical hydration requires harsh acidic conditions (Evans and Kirby 1984). Furthermore, the CrtCs from photosynthetic bacteria act on acyclic carotenoids, whereas the CruFs from non-photosynthetic bacteria catalyze the hydration of monocyclic carotenoids. To our knowledge, no published data exist on the catalytic and structural features that determine hydratase activity and specificity of these two distinct families, nor has the 3D structure been elucidated yet. The mechanism of lycopene hydration, which involves proton attack at C-2 and C-2' and the introduction of the hydroxyl group at C-1 and C-1', was established from <sup>2</sup>H<sub>2</sub>O and H<sub>2</sub><sup>18</sup>O labeling studies with intact cells (Patel et al. 1983; Yeliseev and Kaplan 1997). For a hydration reaction, it is likely to assume that the first step in the reaction is protonation of the alkene, leading to an intermediate carbocation. Quenching of the carbocation by water will lead to the alcohol as product. The protonation of hydrophobic long-chain alkenes has also been described for the enzyme class of cyclases, of which the full mechanism is known (Hammer et al. 2013; Wendt et al. 2000).

The objective of this study was to provide insight into the hydration mechanism of CrtCs. This knowledge is pivotal in order to engineer this promising enzyme class towards, i.e., higher activities, better stability, or widening of its substrate scope. Through multi-sequence alignment of several CrtC homologues, highly conserved amino acids were identified, which could be functionally or structurally important. The corresponding alanine mutants of these amino acids were produced to evaluate their involvement in the hydratase activity. Following the identification of probable catalytically active amino acid residues, the aim was to propose a catalytic mechanism for addition of water catalyzed by CrtC.

## Materials and methods

**In silico analysis** Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990) was used to find and select carotenoid 1,2-hydratase homologues of CrtC from *R. gelatinosus* (RgCrtC) using default settings; e.g., for nucleotide BLAST, the tblastn option was used with the nucleotide database (nt/nr), and for protein BLAST, the blastp option was used with the non-redundant protein database (nr). In order to look for identities/similarities between the CrtC homologues, nucleotide and amino acid sequences were aligned with the BioEdit Sequence Alignment Editor v.7.1.3.0 ([www.mbio.ncsu.edu/bioedit/bioedit.html](http://www.mbio.ncsu.edu/bioedit/bioedit.html)) or ClustalW (Larkin et al. 2007). In addition, protein sequences were subjected to protein functional analysis using a search in the Conserved Domain Database (CDD) (Marchler-Bauer et al. 2011) and Pfam search (Finn et al. 2010), using the standard parameters on the respective websites of these tools. A protein phylogenetic tree was constructed with Phylogeny.fr using the “One Click” program settings (Dereeper et al. 2010; Dereeper et al. 2008). These settings represent a *default* mode which proposes a pipeline already set up to run and connect programs recognized for their accuracy and speed to reconstruct a robust phylogenetic tree from a set of sequences (MUSCLE for multiple alignment, optionally Gblocks for alignment curation, PhyML for phylogeny, and finally, TreeDyn for tree drawing).

**Cloning of carotenoid 1,2-hydratase genes** Plasmids pET15b\_CrtC<sub>Rg</sub> and pET15b\_CrtC<sub>Tr</sub> containing CrtC from *R. gelatinosus* (Rg) and *T. roseopersicina* (Tr), respectively, were constructed in a previous study (Hiseni et al. 2011). Two fosmid genes from metagenomic samples DelRiverFos06H03 (Fos06) and DelRiverFos13D03 (Fos13), respectively, were kindly provided by Dr. Kirchman (Waidner and Kirchman 2005). The cosmid encoding CrtC from *Bradyrhizobium* (Br) was received from Dr. Dreyfus (Giraud et al. 2004). In order to get sufficient DNA material for further studies, the fosmid DNA and cosmid DNA were amplified in *Escherichia coli* TOP10 cells. After DNA isolation with the QIAprep Spin Miniprep Kit (Qiagen) from the cells, sufficient DNA was obtained for further research. The *crtCs* from *Rhodospirillum rubrum* (Rr) and *Rhodopseudomonas palustris* (Rp) were amplified from

genomic DNA. For that, genomic DNA of *R. rubrum* (*Rr*) was kindly provided by Prof. Roberts (NCBI Reference Sequence: NC\_007641.1). *R. palustris* cells (DSM No. 123) were obtained from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen), enriched in appropriate medium according to DSMZ instructions and gDNA isolated using the UltraClean Soil DNA Isolation Kit (Mo Bio). Subsequently, primers were designed for the isolation of all *crtC* genes (Table 1), which carry two restriction sites for subsequent cloning: *NdeI* (forward) and *XhoI* (reverse). For *BrertC*, the *XhoI* site was replaced with *BamHI*, because the *XhoI* site was present in the gene itself. Amplification reactions were done using standard PCR reactions. Using the appropriate enzymes, the fragment was digested, purified, and ligated into the same sites of the pET15b vector and transformed into *E. coli* TOP10 competent cells. The insertion of the *crtC* gene was verified by restriction analysis with the corresponding restriction enzymes (New England Biolabs) and DNA sequencing (BaseClear, Leiden, The Netherlands).

**Single point mutations** Single amino acid exchange within the *crtC* genes of *Rg* and *Tr* was done using the megaprimer PCR method introduced by Kammann et al. (1989) and later modified by Sarkar and Sommer (1990) and Landt et al. (1990). The mismatch primers are listed in Table 2. In the first PCR reaction, performed under standard reaction conditions, the megaprimer was produced using the corresponding forward primer containing the desired base substitution (Table 2) in combination with the reverse primer *Rg\_rv* or *Tr\_rv* (Hiseni et al. 2011). Plasmids pET15b\_*CrtC<sub>Rg</sub>* and pET15b\_*CrtC<sub>Tr</sub>* (Hiseni et al. 2011) were used as the template. The size and purity of the megaprimer was verified by agarose gel electrophoresis. In order to produce the full-length gene, a second PCR reaction was performed with the corresponding megaprimer and *Rg\_fw* or *Tr\_fw* (Hiseni et al. 2011). Subsequent steps were performed as described in the previous section. The presence of the desired mutation was verified by DNA sequencing (BaseClear, Leiden, The Netherlands).

**N-terminally truncated *RgCrtC* and *TrCrtC*** *RgCrtC* and *TrCrtC* lacking the first 45 and 57 amino acids, respectively, were constructed using primers *Rg\_45aa/Rg\_rv* and *Tr\_57aa/Tr\_rv* (Table 3; Hiseni et al. 2011) under standard PCR conditions. Plasmids pET15b\_*CrtC<sub>Rg</sub>* and pET15b\_*CrtC<sub>Tr</sub>* (Hiseni et al. 2011), respectively, were used as the template. Subsequent steps including verification were performed as described in the “Cloning of carotenoid 1,2-hydratase genes” section.

**Recombinant expression of *CrtCs*** *E. coli* BL21 (DE3) was the host for the pET15\_*CrtC* plasmids. Cultures were grown at 37°C in Luria–Bertani broth with 100 µg ml<sup>-1</sup> ampicillin until an OD600 value of 0.6–0.8 was reached. Unless otherwise stated, protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM, followed by cultivation at 25°C overnight. The cells were harvested by centrifugation at 10,000 rpm for 10 min at 4°C (Sorvall), washed once with 50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8.0), and suspended in the same buffer. In case of subsequent purification, 10 mM imidazole was added to the buffer. Crude extract (CE) from cultures >100 ml was prepared by adding 1 mg ml<sup>-1</sup> lysozyme and incubating the cells for 1 h at 4°C, followed by cell disruption at the pressure of 1.5 kBar (Constant Systems; IUL Instruments). For cultures <100 ml, the cells were disrupted by sonication for 2 min while immersed in an ice-water bath using the microtip probe of a sonicator (Branson Sonicor Cell Disruptor) set at 50 % maximal energy. In an effort to reduce the liquid viscosity caused by DNA molecules, 0.1 mg ml<sup>-1</sup> of DNase was added. With the subsequent centrifugation at 10,000 rpm for 20 min at 4°C, cell-free extract (CFE) and pellet were separated. Protein content of the crude extract was determined by BCA assay (Pierce) with bovine serum albumin as the reference protein.

***CrtC* purification** *RgCrtC* and *TrCrtC* “active site” point mutants were purified from the membrane fraction,

**Table 1** Primers used for PCR and subsequent cloning of the genes into the expression vector pET15b

Name	Sequence (5'→3')	Restriction site
DRF06 FW	GGGAGTACC <u>CATATG</u> AGTGATGATGGCCAAC	<i>NdeI</i>
DRF06 RV	ATCCGCTCGAGATAATCTCAAGCCCGCCTCG	<i>XhoI</i>
DRF13 FW	GGGAGTAC <u>CATATG</u> GATGGCGTGTCAGAC	<i>NdeI</i>
DRF13 RV	CCGCTCGAGTAATGCTTAGGGCCACTTGCC	<i>XhoI</i>
Br FW	CGGACATC <u>CATATG</u> TGCCCGCCAG	<i>NdeI</i>
Br RV	ATCCAGGATCCATCGCGTGAACCTCACCACC	<i>BamHI</i>
Rp FW	CGGGACTTCC <u>CATATG</u> TCAGGAGCTGAGTTG	<i>NdeI</i>
Rp RV	ACCGCTCGAGTAACGTTTCAGCGGAACGC	<i>XhoI</i>
Rr FW	GGGAAATTCC <u>CATATG</u> CACCGCCCGGAC	<i>NdeI</i>
Rr RV	GCTCGAGTTCAATTAGCCCTTAACCGCCGC	<i>XhoI</i>

The respective restriction sites are underlined

**Table 2** Primers for site-directed mutagenesis

	Amino acid exchange	Sequence (5'→3')
<i>R. gelatinosus</i>	H239A	AGCGGCGGACGCGCTCGCTG
	W241A	CATCGCGCGGGCCGATCG
	H264A	CTGGAGCGGCGCCGCTACC
	Y266A	GCCACGCCGCCCTCGACT
	D268A	CGCTACCTCGCCTCGAACGAAG
<i>T. roseopersicina</i>	H237A	GATCCGGCGGAACGCGCAGTCTGGTGG
	W239A	CGCCATGTGCGGTGGCCGATC
	H262A	GCTGGAGCGGCGCTGGCTAT
	D266A	CATGGCTATCTCGCCTCAAA
	S58V	GCGTCCGTCGTCGCGCAGCA
	S58Q	GCGTCC <u>AGG</u> TCGCGCAGCA

Mismatch points are underlined

while the *Tr* “processing” mutants (S58V, S58Q) were purified from the CFE. The membrane fraction was obtained after the centrifugation of the CFE for 4 h at 13, 200 rpm and 8°C. The membranes were homogenized by ca. 20 passages through a 25-G needle. Ni-NTA Sepharose HP (GE Healthcare) (previously equilibrated in 50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer, pH 8.0, 300 mM NaCl, 10 mM imidazole) was added to the CFE or membrane sample. The mixtures were incubated for 1 h at RT, loaded into a polypropylene tube with porous disc (GE Healthcare), and washed three times with washing buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer, pH 8.0, 300 mM NaCl, 75 mM imidazole). The CrtC protein was eluted from the column with elution buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub> buffer, pH 8.0, 300 mM NaCl, 1 M imidazole). Enzyme fractions were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (10 % Bis-Tris; Bio-Rad) and visualized by staining with SimplyBlue SafeStain (Invitrogen).

**Determination of enzyme activity** Enzymatic activities were determined with CE using lycopene as the substrate according to the method described earlier (Hiseni et al. 2011), with few modifications. The assay was carried out with 50 µl CE and 20 µM substrate, and 10 mg ml<sup>-1</sup> L-α-phosphatidylcholine, in a reaction volume of 200 µl. Prior to the analysis, acetonitrile was added to the reaction mixture in a ratio of 60:40 (ACN/H<sub>2</sub>O), the mixtures were shaken vigorously for 1 min, and

solids were removed by centrifugation for 1 min at 13, 200 rpm. Separation of the reaction products was performed with a Merck 4.6×50 mm Chromolith® SpeedROD RP-18e column with ACN/H<sub>2</sub>O (60:40, v/v) as the eluent.

## Results

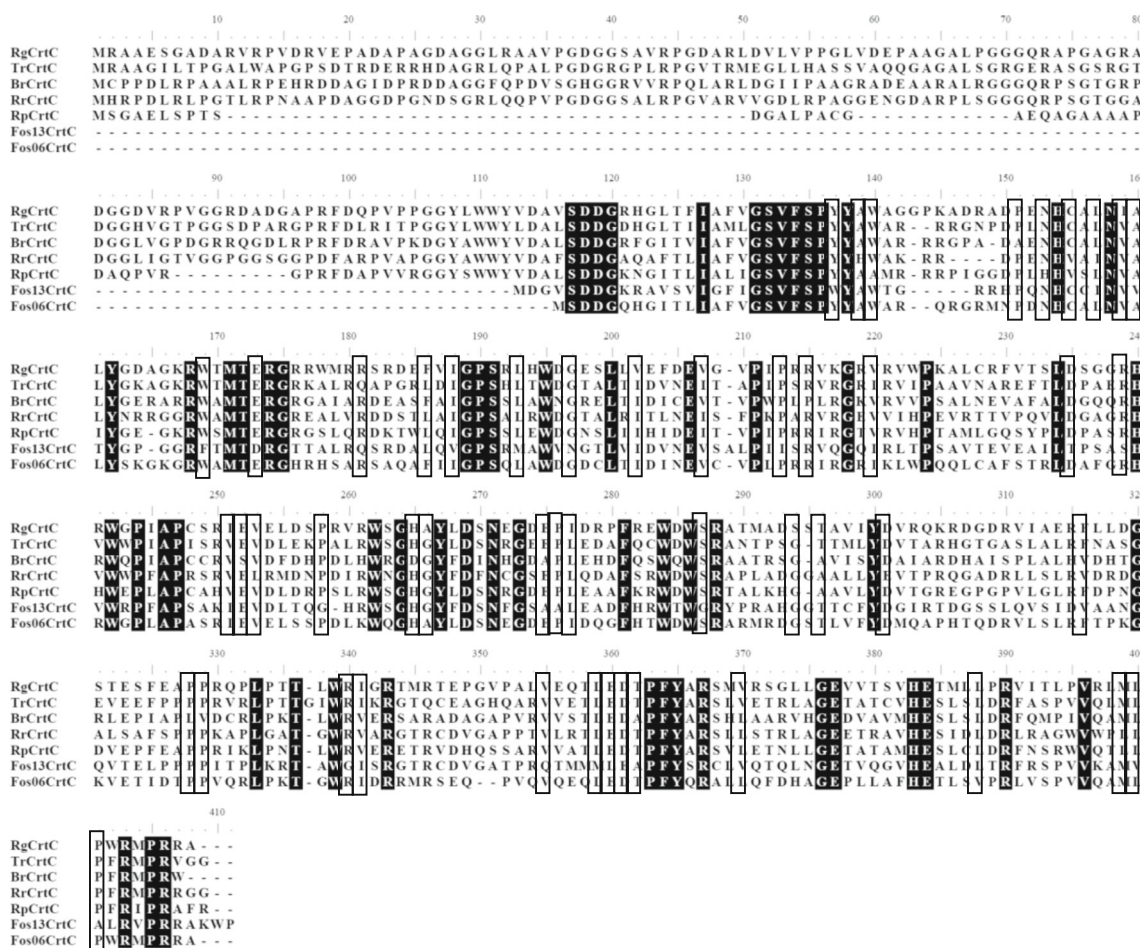
**Comparative in silico analysis of *crtC* genes** The *Rg-crtC* nucleotide sequence was subjected to a BLAST search in order to identify sequence similarity in different databases. One hundred forty-five hits were identified, of which 119 were representatives of Proteobacteria. BLASTing the protein sequence of *RgCrtC* resulted in 100 CrtC(-like) sequences, which were aligned in order to investigate if there are any conserved group clusters present (see Electronic Supplementary Material Fig. S1). Indeed, they showed highly conserved regions (Fig. 2, for simplification, the seven sequences used in this study were selected for visualization) distributed along the sequence ranging from amino acid residues ~170 to ~405 (*Rg* numbering). Interestingly, the N-terminal part of the sequence does not contain any conserved amino acids and shows a lot of variation in length, indicating that this region is probably not necessary for CrtC activity. Additionally, we found some blocks of highly conserved residues throughout the C-terminal part of the genes and a cluster of totally conserved residues in the middle of the genes (amino acids 240–280 in *Rg* numbering).

Residues involved in the catalysis tend to be highly conserved in a set of homologous proteins that exhibit the same reaction. On the other hand, sequence insertion and sections of low sequence similarity tend to occur in the less important loop regions (Zvelebil et al. 1987). The conserved blocks in CrtC homologues could indicate that these regions contain the amino acid residues most important for the hydratase activity, specifically those involved in catalysis and substrate binding.

**Table 3** Primers used for construction of truncated CrtCs

Name	Sequence (5'→3')
Rg_45aa	AGTACCATATGGGCGACGCACGGCTGG
Tr_57aa	AGTACCATATGTCCGTCGCGCAGCAAGG

The *NdeI* restriction sites are underlined



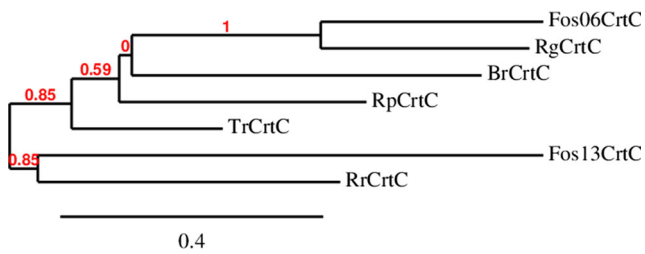
**Fig. 2** Multiple sequence alignment of CrtC protein sequences probed in this study showing conserved amino acids. Identical amino acids are highlighted in *black*. Positions with only two different amino acids are surrounded by *boxes*

From a catalytic point of view, specific amino acids are commonly involved as active residues in acid–base-type catalyzed reactions in the active sites of enzymes (Bartlett et al. 2002). The amino acids aspartic acid (D) or glutamic acid (E) are usually the catalytic acid or base, while tyrosine (Y), tryptophan (W), and histidine (H) typically function as the other part of the charge relay pair (Puthan Veetil et al. 2012). These amino acids are, therefore, the most probable candidates for the catalytic hydration. Four Trp residues, three Tyr residues, and one of each His and Asp residues are fully conserved over the 100 CrtCs (Supplemental Fig. S1). The totally conserved Asp268 (*Rg* numbering) seems to be the most probable candidate for the acid catalysis since it is situated in the middle of the highly conserved region. Furthermore, a fully conserved Tyr residue and a His residue are within close distance, which is important for the contact with the substrate. Fully conserved His and Trp residues are approximately 30 amino acids away but could be in close proximity upon folding of the protein.

**Production of recombinant wild type and mutant CrtCs and enzymatic activity**

Six potential CrtCs were selected for

expression and activity studies based on sequence identity with *RgCrtC* and availability of the corresponding gene constructs. They originate from all three Proteobacteria subclasses (Alphaproteobacteria, Betaproteobacteria, and Gammaproteobacteria) with two additional constructs originating from metagenomic samples from the Delaware River (USA). Figure 3 displays the phylogenetic analysis constructed with protein sequences of the selected CrtC homologues. *TrCrtC* shows the closest relationship to *RgCrtC*, followed by *BrCrtC* (55 and 47 % sequence identity, respectively). The combined results of Pfam search and Conserved Domain Database search showed that all seven CrtCs belong to the PF07143 family consisting of several purple photosynthetic bacterial hydroxynurosporene synthase (CrtC) proteins. Six out of the seven selected CrtCs could be overexpressed from pET15b in *E. coli* (Fig. 4). Bands with apparent molecular weight of 32 kDa (Fos13CrtC), 38 kDa (*RpCrtC*), and 44 kDa (*RrCrtC*, *BrCrtC*, *TrCrtC*, and *RgCrtC*) were visualized on SDS-PAGE and were consistent with the values calculated from the deduced amino acid sequences. *TrCrtC* shows two protein bands of 44 and 38 kDa as seen before



**Fig. 3** Rooted phylogenetic tree showing the evolutionary relationship between the selected carotenoid 1,2-hydratases. The scale bar represents 0.4 substitutions per amino acid site. *TrCrtC* *Thiocapsa roseopersicina* (GI 31621263), *BrCrtC* *Bradyrhizobium* sp. BTAi1 (GI 146403799), *Fos06CrtC* uncultured Proteobacterium DelRiverFos06H03 (GI 61653228), *Fos13CrtC* uncultured Proteobacterium DelRiverFos13D03 (GI 61653190), *RpCrtC* *Rhodospseudomonas palustris* (GI 115515977), *RrCrtC* *Rhodospirillum rubrum* (GI 83574254), *RgCrtC* *Rubrivivax gelatinosus* (GI 29893477)

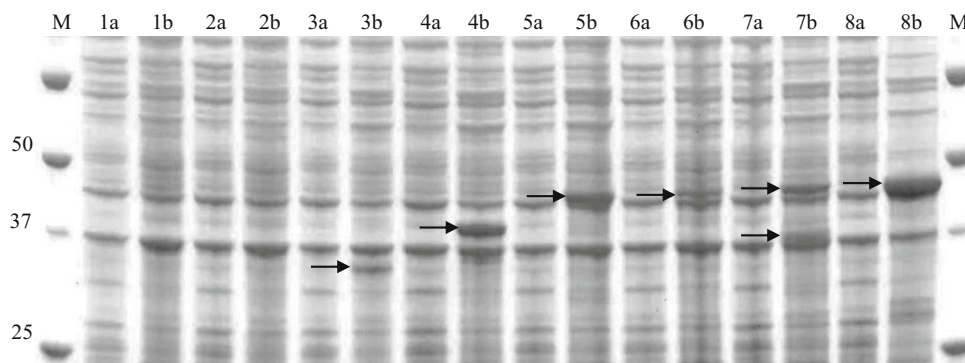
(Hiseni et al. 2011). No expression band could be identified for Fos06CrtC. Although relatively good expression was achieved for most of the CrtCs, only two were active with lycopene as the substrate in the standard enzymatic assay, i.e., *RgCrtC* and *TrCrtC* (data not shown).

The amino acid residues that might be important for catalysis (*vide supra*) were substituted by the amino acid alanine in both active enzymes. In addition, truncated (*TrCrtC* and *RgCrtC*) and N-terminal point mutants (*TrCrtC*) were constructed and analyzed to confirm the importance of the N-terminal part of CrtC for the catalytic activity. The activity of the truncated versions was fully retained. Despite the still unknown reason for the truncation, we were able to identify the cleavage site between S57 and S58 by MS analysis. In order to exclude that this truncation only takes place in recombinant expression, the S58 position was modified by substitutions with valine or glutamine (see Electronic Supplementary Materials, including Fig. S2).

All mutants (Table 2) were successfully cloned and expressed in *E. coli* BL21 (Fig. 5a). However, clear

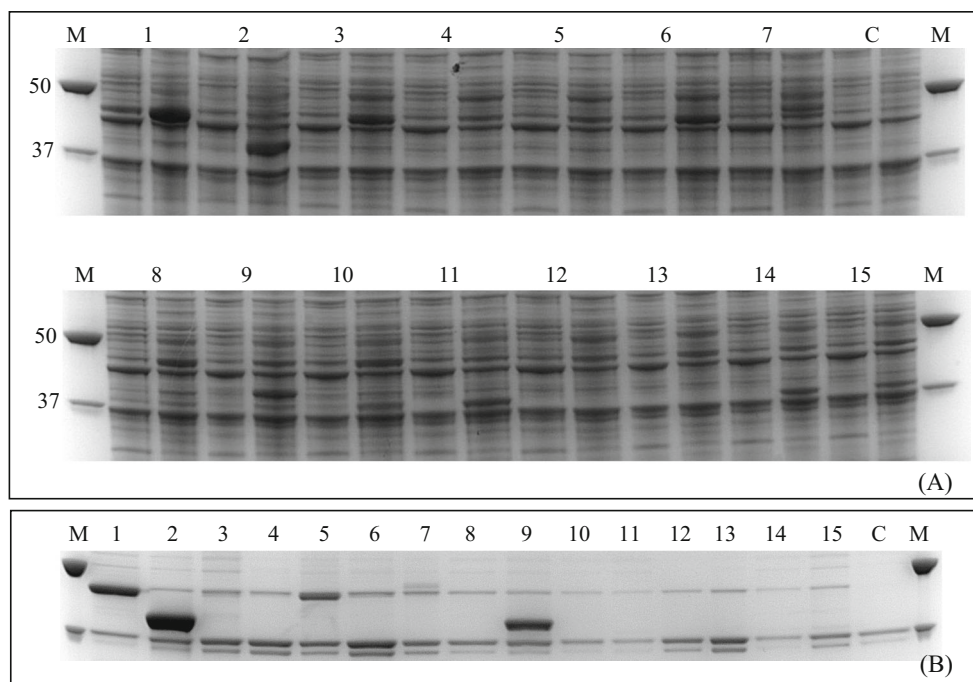
differences in expression levels were observed. While the removal of the N-terminus resulted in an increased expression level, all point mutations negatively influenced the expression of the protein. In order to ensure that CrtC was present, all mutants were purified from the membrane fraction. As can be seen in Fig. 5b, all mutants could be purified by chromatography over a Ni column binding the His tag and showed a band at 38 or 44 kDa, which was absent in the negative control sample (pET15b). In the case of *TrCrtC* constructs, only very weak protein bands are detected after the purification (Fig. 5b, lanes 8 and 10–15). The purification usually has to be performed as soon as it is expressed (before the cleavage of the N-terminal part including the His tag), which was not the case here.

Next to the analysis of the expression levels, the activities of all constructed mutants were measured with lycopene as the substrate (Fig. 6). As the expression and purification levels were very low for some of the mutants and the activity of CrtC, in general, is very low, crude extracts were used for the activity assays. Consequently, the results cannot be quantitatively compared. However, in combination with the expression levels as shown in Fig. 5a, indicative conclusions can be drawn. When looking at the results from the alanine mutants, it appears that four key residues were identified, which have a potentially important role in the hydration mechanism. By replacing each of the amino acids H239, W241, Y266, and D268 individually by an alanine in *RgCrtC*, the activity is completely destroyed. The same mutations of the corresponding amino acids in *TrCrtC*, i.e., H237, W239, and D266, also resulted in CrtC inactivation. Unfortunately, the mutagenesis of Y264 in *TrCrtC* was not successful and, therefore, could not be included in this study. However, based on all the results, one could expect that the mutation of Y264 in *TrCrtC* would lead to inactivation, as seen in *RgCrtC*. On the other hand, the less conserved H264 in *RgCrtC* and the corresponding histidine residue in *TrCrtC* (H262) seem not to have any functional role. The mutants fully retained activity and



**Fig. 4** SDS-PAGE (10 %) analysis of CrtC expression in *E. coli* BL21. *M* precision plus protein standard. The first lane (a) of each sample shows cells before induction with 0.6 mM IPTG, and the second lane (b) shows cells after 4 h of expression at 37 °C. 1 pET15b control, 2 Fos06CrtC

(32 kDa), 3 Fos13CrtC (32 kDa), 4 *RpCrtC* (38 kDa), 5 *RrCrtC* (44 kDa), 6 *BrCrtC* (44 kDa), 7 *TrCrtC* (44 kDa), 8 *RgCrtC* (44 kDa). The indicated molecular weights are deduced from amino acid sequences. CrtC expression bands are indicated by arrows



**Fig. 5** SDS-PAGE (10 %) analysis of expression (a) and IMAC purification from membrane (b) of CrtCs from *R. gelatinosus* (*RgCrtC*) (lanes 1–7) and *T. roseopersicina* (*TrCrtC*) (lanes 8–15) wild type and mutants. *M* precision plus protein standard. *C* pET15b control. **a** The first lane of each sample shows cells before induction with 0.1 mM IPTG, and

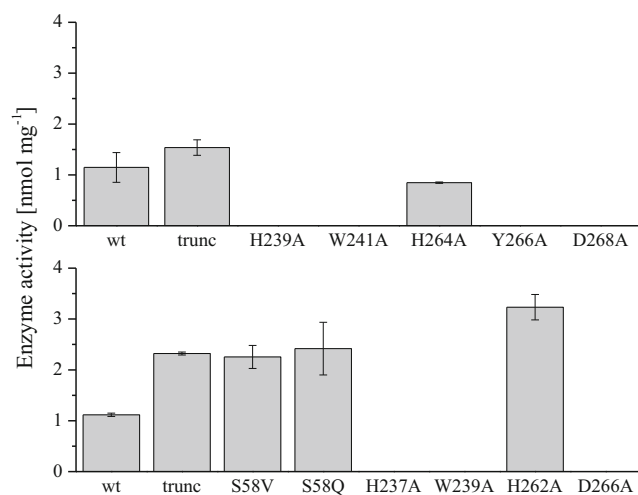
the second lane shows cells after overnight expression at 25 °C. 1 *RgCrtC* wild type, 2 *RgCrtC* truncated, 3 *RgCrtC* H239A, 4 *RgCrtC* W241A, 5 *RgCrtC* H264A, 6 *RgCrtC* Y266A, 7 *RgCrtC* D268A, 8 *TrCrtC* wild type, 9 *TrCrtC* truncated, 10 *TrCrtC* S58V, 11 *TrCrtC* S58Q, 12 *TrCrtC* H237A, 13 *TrCrtC* W239A, 14 *TrCrtC* H262A, 15 *TrCrtC* D266A

even showed slightly increased activity when the expression levels were considered. For instance, the truncated *TrCrtC* and H262A mutant showed almost the same level of expression (Fig. 5a, lanes 9 and 14) but the activity of H262A mutant was ~1.3-fold higher (Fig. 6). The same was observed for *RgCrtC*, where the expression of the wild type is much higher than that

of the mutant H264A, but both showed approximately the same activity.

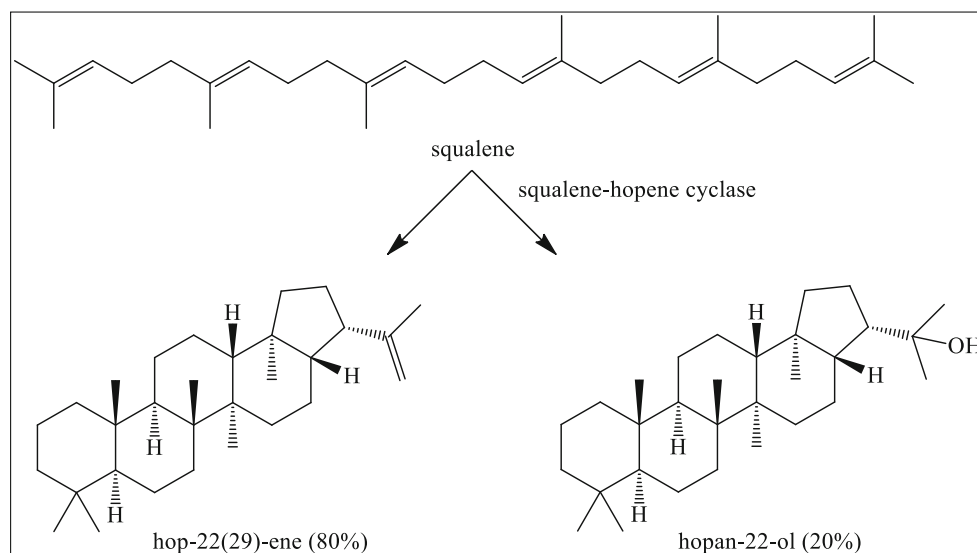
## Discussion

The main purpose of this study was to get more insight into the hydration mechanism of carotenoid 1,2-hydratases. First, the distribution of these enzymes was assessed by subjecting the *Rg-crtC* nucleotide sequence to a BLAST search. Although *R. gelatinosus* belongs to the Betaproteobacteria, more than 69 % of the identified 119 hits were from Alphaproteobacteria and only 11 % from Betaproteobacteria. Similarly, Igarashi et al. (2001) observed that most of the photosynthesis gene products from *R. gelatinosus* showed high sequence identities to the gene products of *R. palustris*, an Alphaproteobacteria member. They explain this occurrence as horizontal transfer of the photosynthesis gene clusters from an ancestral species belonging to the Alphaproteobacteria to that of the Betaproteobacteria, which might also have happened to the CrtC genes. The identified CrtC sequences were aligned in order to discover conserved group clusters. In total, 33 amino acids were found to be fully conserved. No conserved residues were identified in the N-terminal part of the sequence (amino acids 1–125 in *Rg* numbering), which is in agreement with our earlier hypothesis that this region is probably not necessary for CrtC activity (Hiseni et al. 2011). This was also confirmed by



**Fig. 6** Enzymatic activity of wild type (*wt*) and mutant CrtC from *R. gelatinosus* (upper) and *T. roseopersicina* (lower). Extracts from *E. coli* cells expressing the respective enzymes were assayed with 20 μM lycopene in 50 mM Na<sub>2</sub>HPO<sub>4</sub> sodium phosphate (pH 8.0) at 28 °C overnight. *trunc* variants with missing N-terminal residues 1–45 (*RgCrtC*) and 1–57 (*TrCrtC*)

**Fig. 7** Enzyme-catalyzed cyclisation of squalene to hopene and hopanol



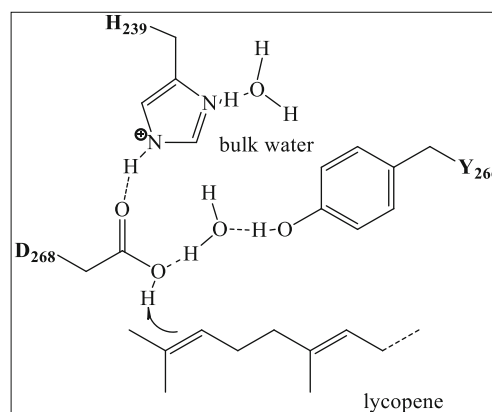
the activity of the truncated variants. Furthermore, this also accounts for the absence of this part and, thus, the shorter DNA sequences for many CrtCs, including Fos06, Fos13, and partly *RpCrtC*, when compared to *RgCrtC* (Fig. 2, Supplemental Fig. S1). This part is, hence, not involved in the catalytic mechanism.

Six out of the seven selected CrtCs could be overexpressed in *E. coli* (Fig. 4). The fact that all CrtCs share highly conserved regions in the amino acid sequence indicates that they are performing the same or similar biochemistry. However, no activity whatsoever could be detected for four out of six CrtCs in the standard lycopene hydration assay. At this point, it is unclear whether this is due to reasons of low activity in the cell extract and/or substrate specificity. Since CrtCs are active in different parts of the carotenoid pathway, substrate specificities can differ depending on the carotenoid produced by the organism where the enzyme originated (Kovács et al. 2003).

We have stated earlier that the hydration can take place through an acid–base-type catalysis, resulting in the identification of five possible key residues. The identified key residues H239, W241, Y266, and D268 in *RgCrtC* and the corresponding residues in *TrCrtC* were probed by generating alanine point mutants thereof. The absence of activity upon individual substitution leads to the hypothesis that they are involved in the hydration catalysis. Furthermore, they are all in close distance to each other on the protein chain (Fig. 2). The fact that there is only one other residue between H239 and W241 or Y266 and D268 is consistent with residues in space pointing in the same direction in a beta strand or loop (one residue pointing upward, the next downward). In order to investigate how these newly identified key residues could be involved in the catalytic hydration reaction, we built a 3D structure of *TrCrtC* by homology modeling. The closest known 3D structure, the putative AttH protein from *Nitrosomonas europaea*, showed only a sequence identity of

17 % to the CrtC (see Electronic Supplementary Material Figs. S3, S4, and S5). This is not enough for a reliable model. However, this model does indicate the possibility of the presence of such an active site.

These four residues, which are conserved throughout the whole CrtC family, are also found in the active site of squalene–hopene cyclase (SHC) (Wendt et al. 2000). SHC catalyzes the cyclization reaction of squalene to hopene as a major product (Fig. 7). Hopanol is also formed to a minor extent. The proposed mechanism for cyclases is proton-triggered polycyclization, whereby the intermediate carbocation is stabilized by aromatic amino acids. Next to the stabilization role of the aromatic amino acids, they also create a hydrophobic environment in order to prevent quenching of the cation by water. The cyclization cascade is terminated by a well-positioned enzymatic base. The formation of the alcohol side product suggests significant water accessibility at the termination region of the active site. The acidic residue aspartate (D376), which



**Fig. 8** Proposed mechanism for the initial protonation during lycopene hydration



is located in the center of the active site in SHC, is the likely general acid responsible for protonating the C3 atom of the squalene substrate (Wendt et al. 2000). The acidity of D376 is enhanced by a connection to the side chain of Y495 through a water molecule. The tyrosine positions the proton on aspartic acid into the anti-orientation, turning it more acidic. Carboxylic acid protons in anti-orientation have been estimated to be  $10^4$  times more acidic than the biologically more relevant syn-oriented protons (Gandour 1981).

Because of the probable similarity of the initial protonation reactions of squalene and lycopene, we assume that the residues involved in catalysis will be alike in SHC and CrtC. Therefore, we propose the following mechanism for RgCrtC. D268 is the catalytic acid that initiates the hydration of lycopene (Fig. 8). Upon diffusion of lycopene into the active site, the C2 atom of the substrate needs to be positioned near the proton of D268 that putatively will be added onto the substrate. In order to enhance the acidity of the catalytic D268 for olefin protonation, the amino acid is directly bonded to H239 and to Y266 through an ordered water molecule, similar to what has been proposed for SHC (Wendt et al. 2000). This hypothesis is supported by the results that mutation of one of these three amino acids leads to inactivation of enzymatic activity. In contrast to SHC, where premature quenching of the cationic intermediate by water or nucleophiles is prevented by well-positioned aromatic amino acids, a water molecule is added to lycopene to yield the desired hydroxylated lycopene derivative. This suggests that the active site of CrtC has more water molecules present, so that the interaction between the substrate and solvent water molecules is more significant. The aromatic amino acid Trp266 might be involved in the correct positioning of the hydrophobic substrate and the stabilization of the intermediate carbocation.

In conclusion, overall results indicate that residues H239, W241, Y266, and D268 in RgCrtC are probably involved in an acid–base-type hydration. The absence of activity upon individual substitution of these residues by an alanine agrees with a role in the catalytic cycle. We hypothesize that they are involved in the initial protonation, which would be followed by quenching of the carbocation by a water molecule, resulting in the hydration product. From our findings, it becomes clear that the complete structure of a CrtC, through crystallization studies, will be pivotal to really unravel the mechanism for this intriguing enzyme. Nevertheless, the results of this study produce for the first time a workable hypothesis for the catalytic mechanism of carotenoid 1,2-hydratase and open the field for the engineering of this enzyme towards industrially relevant mutants.

**Acknowledgments** We thank Dr. Jaap Jongejan (Delft University of Technology) for his advice on finding the potentially important residues in carotenoid 1,2-hydratases.

**Compliance with ethical standards** This project was financially supported by The Netherlands Ministry of Economic Affairs and the B-Basic partner organizations (<http://www.b-basic.nl>) through B-Basic, a public–private NWO-ACTS programme [Advanced Chemical Technologies for Sustainability (ACTS)].

**Conflict of interest** The authors declare that they have no competing interests.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410. doi:10.1016/S0022-2836(05)80360-2
- Armstrong GA (1997) Genetics of eubacterial carotenoid biosynthesis: a colorful tale. *Annu Rev Microbiol* 51(1):629–659. doi:10.1146/annurev.micro.51.1.629
- Armstrong GA, Hearst JE (1996) Carotenoids 2: genetics and molecular biology of carotenoid pigment biosynthesis. *FASEB J* 10(2):228–237
- Bartlett GJ, Porter CT, Borkakoti N, Thomson JM (2002) Analysis of catalytic residues in enzyme active sites. *J Mol Biol* 324(1):105–121. doi:10.1016/S0022-2836(02)01036-7
- Cazzonelli CI (2011) Carotenoids in nature: insights from plants and beyond. *Funct Plant Biol* 38(11):833–847. doi:10.1071/FP11192
- Dereeper A, Audic S, Claverie J-M, Blanc G (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol Biol* 10(1):1–6. doi:10.1186/1471-2148-10-8
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, Claverie J-M, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36(suppl 2):W465–W469. doi:10.1093/nar/gkn180
- Evans CM, Kirby AJ (1984) A model for olefin hydration: intramolecular nucleophilic addition of phenolate oxygen to the unactivated double bond. *J Chem Soc Perkin Trans 2*(7):1259–1267. doi:10.1039/P29840001259
- Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A (2010) The Pfam protein families database. *Nucleic Acids Res* 38(suppl 1):D211–D222. doi:10.1093/nar/gkp985
- Gandour RD (1981) On the importance of orientation in general base catalysis by carboxylate. *Bioorg Chem* 10(2):169–176. doi:10.1016/0045-2068(81)90020-1
- Giraud E, Hannibal L, Fardoux J, Jaubert M, Jourand P, Dreyfus B, Sturgis JN, Verméglio A (2004) Two distinct *crt* gene clusters for two different functional classes of carotenoid in *Bradyrhizobium*. *J Biol Chem* 279(15):15076–15083. doi:10.1074/jbc.M312113200
- Hammer SC, Syrén P-O, Seitz M, Nestl BM, Hauer B (2013) Squalene hopene cyclases: highly promiscuous and evolvable catalysts for

- stereoselective CC and CX bond formation. *Curr Opin Chem Biol* 17(2):293–300. doi:10.1016/j.cbpa.2013.01.016
- Hiseni A, Arends IWCE, Otten LG (2011) Biochemical characterization of the carotenoid 1,2-hydratases (CrtC) from *Rubrivivax gelatinosus* and *Thiocapsa roseopersicina*. *Appl Microbiol Biotechnol* 91(4):1029–1036. doi:10.1007/s00253-011-3324-1
- Hiseni A, Arends IWCE, Otten LG (2015) New cofactor-independent hydration biocatalysts: structural, biochemical, and biocatalytic characteristics of carotenoid and oleate hydratases. *ChemCatChem* 7(1):29–37. doi:10.1002/cctc.201402511
- Igarashi N, Harada J, Nagashima S, Matsuura K, Shimada K, Nagashima KVP (2001) Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria. *J Mol Evol* 52(4):333–341. doi:10.1007/s002390010163
- Jensen SL, Cohen-Bazire G, Stanier RY (1961) Biosynthesis of carotenoids in purple bacteria: a re-evaluation based on considerations of chemical structure. *Nature* 192(4808):1168–1172. doi:10.1038/1921168a0
- Jin J, Hanefeld U (2011) The selective addition of water to C=C bonds; enzymes are the best chemists. *Chem Commun* 47(9):2502–2510. doi:10.1039/C0CC04153J
- Kammann M, Laufs J, Schell J, Gronenbom B (1989) Rapid insertional mutagenesis of DNA by polymerase chain reaction (PCR). *Nucleic Acids Res* 17(13):5404. doi:10.1093/nar/17.13.5404
- Kourist R, Domínguez de María P, Bornscheuer UT (2008) Enzymatic synthesis of optically active tertiary alcohols: expanding the biocatalysis toolbox. *ChemBioChem* 9(4):491–498. doi:10.1002/cbic.200700688
- Kovács ÁT, Rákhely G, Kovács KL (2003) Genes involved in the biosynthesis of photosynthetic pigments in the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina*. *Appl Environ Microbiol* 69(6):3093–3102. doi:10.1128/aem.69.6.3093-3102.2003
- Landt O, Grunert H-P, Hahn U (1990) A general method for rapid site-directed mutagenesis using the polymerase chain reaction. *Gene* 96(1):125–128. doi:10.1016/0378-1119(90)90351-q
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948. doi:10.1093/bioinformatics/btm404
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Lu F, Marchler GH, Mullokandov M, Omelchenko MV, Robertson CL, Song JS, Thanki N, Yamashita RA, Zhang D, Zhang N, Zheng C, Bryant SH (2011) CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res* 39(Suppl 1):D225–D229. doi:10.1093/nar/gkq1189
- Patel NJ, Britton G, Goodwin TW (1983) Use of deuterium labelling from deuterium oxide to demonstrate carotenoid transformations in photosynthetic bacteria. *Biochim Biophys Acta, Gen Subj* 760(1):92–96. doi:10.1016/0304-4165(83)90128-9
- Puthan Veetil V, Fibriansah G, Raj H, Thunnissen A-MWH, Poelarends GJ (2012) Aspartase/fumarase superfamily: a common catalytic strategy involving general base-catalyzed formation of a highly stabilized aci-carboxylate intermediate. *Biochemistry* 51(21):4237–4243. doi:10.1021/bi300430j
- Sarkar G, Sommer SS (1990) The ‘megaprimer’ method of site-directed mutagenesis. *BioTechniques* 8(4):404–407
- Sun Z, Shen S, Wang C, Wang H, Hu Y, Jiao J, Ma T, Tian B, Hua Y (2009) A novel carotenoid 1,2-hydratase (CruF) from two species of the non-photosynthetic bacterium *Deinococcus*. *Microbiology* 155(8):2775–2783. doi:10.1099/mic.0.027623-0
- Umeno D, Tobias AV, Arnold FH (2005) Diversifying carotenoid biosynthetic pathways by directed evolution. *Microbiol Mol Biol Rev* 69(1):51–78. doi:10.1128/mmbr.69.1.51-78.2005
- Waidner LA, Kirchman DL (2005) Aerobic anoxygenic photosynthesis genes and operons in uncultured bacteria in the Delaware River. *Environ Microbiol* 7(12):1896–1908. doi:10.1111/j.1462-2920.2005.00883.x
- Wendt KU, Schulz GE, Corey EJ, Liu DR (2000) Enzyme mechanisms for polycyclic triterpene formation. *Angew Chem Int Ed* 39(16):2812–2833. doi:10.1002/1521-3773(20000818)39:16<2812::AID-ANIE2812>3.0.CO;2-#
- Yeliseev AA, Kaplan S (1997) Anaerobic carotenoid biosynthesis in *Rhodobacter sphaeroides* 2.4.1: H<sub>2</sub>O is a source of oxygen for the 1-methoxy group of spheroidene but not for the 2-oxo group of spheroidenone. *FEBS Lett* 403(1):10–14. doi:10.1016/S0014-5793(97)00006-9
- Zvelebil MJ, Barton GJ, Taylor WR, Sternberg MJE (1987) Prediction of protein secondary structure and active sites using the alignment of homologous sequences. *J Mol Biol* 195(4):957–961. doi:10.1016/0022-2836(87)90501-8